## Glossary

# IT Automation with Python

## **Terms and definitions from all courses**

## A

**A/B testing:** A way to compare two versions of something to find out which version performs better

**Absolute path:** A full path to the resource in the file system

**Activity monitor:** Mac OS tool that shows what's using the most CPU, memory, energy, disk, or network

**Alteration:** RegEx that matches any one of the alternatives separated by the pipe symbol

**API endpoint:** The part of the program that listens on the network for API calls

**Application Programming Interface (API) key:** This is an authentication token that calls an API, which is then called to identify the person, programmer, or program trying to access a website

**Artifact:** A byproduct of the software development process that can be accessed and used, an item produced during programming

**Automation:** The process of replacing a manual step with one that happens automatically, essential tool for keeping up with the infrastructure needs of a growing business

**Automatic scaling:** This service uses metrics to automatically increase or decrease the capacity of the system

**Automatic testing:** A process where software checks itself for errors and confirms that it works correctly

**Autoscaling:** Allows the service to increase or reduce capacity as needed, while the service owner only pays for the cost of the machines that are in use at any given time

## B

**Backreference:** This is applied when using re.sub() to substitute the value of a capture group into the output

**Bandwidth:** How much data can be sent or received in a second

**Bash:** The most commonly used shell on Linux

**Bash script:** A script that contains multiple commands

**Binary search:** A search algorithm used to find a specific item in a sorted list or array by repeatedly dividing the search space in half until the desired item is found or determined to be absent

**Bisecting:** Dividing in two, also a Git command

**Black-box tests:** A test where there is an awareness of what the program is supposed to do but not how it does it

**Branch:** A pointer to a particular commit, representing an independent line of development in a project

**Break:** A way to exit out of a loop before the loop's condition is false

**Breakpoints:** Debugging features that lets code run until a certain line of code is executed

**Built-in functions:** Functions that exist within Python and can be called directly

## C

**Cache:** This stores data in a form that's faster to access than its original form

**Capacity:** How much the service can deliver

**Centralized logs collection:** This means there's a special server that gathers all the logs from all the servers, or even all computers in the network

**Character classes:** These are written inside square brackets and let us list the characters we want to match inside of those brackets

**Character ranges:** Ranges used to match a single character against a set of possibilities

**CI/CD:** The name for the entire continuous integration and continuous deployment system

**Client-side scripting language:** Primarily for web programming; the scripts are transferred from a web server to the end-user’s internet browser, then executed in the browser

**Code editors:** Tools to provide features, including syntax highlighting, automatic indentation, error checking, and autocompletion

**Code reviews:** The deliberate and methodical gathering of other programmers to examine each other's code for errors to increase the code quality and reduces the amount of bugs

**Cold data:** Accessed infrequently and stored in cold storage

**Comma separated values (CSV):** A very common data format used to store data as segment of text separated by commas

**Command line arguments:** Inputs provided to a program when running it from the command line

**Comments:** Notes to yourself and/or other programmers to make the purpose of the code clear

**Communications lead:** The lead person who needs to receive timely important communication updates

**Commit:** A command to make edits to multiple files and treat that collection of edits as a single change

**Commit files:** A stage where the changes made to files are safely stored in a snapshot in the Git directory

**Commit ID:** An identifier next to the word commit in the log

**Commit message:** A summary and description with contextual information on the parts of the code or configuration of the commit change

**Compiled or compiled language:** Supportable code that can be executed on different platforms

**Computer program:** A step-by-step list of instructions that a computer follows to reach an intended goal

**Computer protocols:** Guidelines published as open standards so that any given protocol can be implemented in various products

**Configuration management:** Automation technique that manages the configuration of computers at scale

**Containers:** Applications that are packaged together with their configuration and dependencies

**Container registry:** A storage location for container images, organized for efficient access

**Container repository:** A container registry that manages container images

**Content Delivery Networks (CDN):** A network of physical hosts that are geographically located as close to the end users as possible

**Continuous delivery:** Any changes to the software are tested and then deployed to users and servers as soon as they are verified

**Continuous deployment:** Automates the deployment of code to production

**Continuous integration:** Constantly adding updates and improvements to software

**Continuous deployment (CD):** New code is deployed often after it has been automatically built and tested

**Continuous integration (CI):** A system that will automatically build and test our code every time there's a change

**Control statements:** Programming constructs that direct the flow of execution of a program by allowing you to make decisions, repeat actions, or choose between different code paths based on specific conditions

**Core files:** Files that store all the information related to the crash to debug the issue

**Cut:** A command that can split and take only bits of each line using spaces

## D

**Data types:** Classes of data (e.g., string, int, float, Boolean, etc.), which include the properties and behaviors of instances of the data type (variables)

**Data serialization:** The process of taking an in-memory data structure, like a Python object, and turning it into something that can be stored on disk or transmitted across a network

**Debuggers:** Tools that follow the code line by line, inspect changes in variable assignments, interrupt the program when a specific condition is met, and more

**Debugging:** The process of identifying, analyzing, and removing bugs in the actual code of a system in the application

**Decorator:** Used in Python to add extra behavior to functions without having to modify the code

**DevOps:** Describes the steps of the software development lifecycle beyond writing code, the union between the development team and the operations team

**DevSecOps:** Adding security testing and protection to the software development lifecycle

**Dialects:** Rules that define how a CSV file is structured

**Dictionaries:** A data type used to organize elements into collections, taking the form of pairs of keys and values

**Diff:** A command to find the differences between two files

**Distributed:** Each developer has a copy of the whole repository on their local machine

**Distributed systems:** Also referred to as distributed computing or distributed databases, utilize different nodes to interact and synchronize over a shared network

**Disk image:** A snapshot of a virtual machine’s disk at a given point in time

**Docstrings:** Documentation that lives alongside the code

**Docker:** An open-source tool used to build, deploy, run, update, and manage containers

**Domain-Specific Language (DSL):** A programming language that's more limited in scope

**DNS zone file:** A configuration file that specifies the mappings between IP addresses and host names in your network

## E

**Edge cases:** Inputs to code that produce unexpected results, found at the extreme ends of the ranges of input

**Environment variables:** Settings and data stored outside a program that can be accessed by it to alter how the program behaves in a particular environment

**Ephemeral storage:** Storage used for instances that are temporary and only need to keep local data while they’re running

**Error budgets:** Represented as the maximum amount of time that a program is able to fail without violating an agreement

**Executor:** This is the process that's in charge of distributing the work among the different workers

**Exhausted:** When resources are used completely and programs are getting blocked by not having more access to those resources

**Expensive actions:** Actions that can take a long time to complete

**Explicit conversion:** This occurs when code is written to manually convert one data type to another using a data type conversion function

**Expression:** A combination of numbers, symbols, or other values that produce a result when evaluated

## F

**Facts:** Variables that represent the characteristics of the system

**Fast-forward merge:** A merge when all the commits in the checked out branch are also in the branch that's being merged

**File systems:** Methods and structures used to organize and control how data is stored and accessed

**Fix up:** The decision to discard commit messages for that commit

**Flask:** A Python library that makes it easier to create web applications and REST web services

**For loop:** This executes a block of code for a specified number of iterations or over a collection of items

**Forking:** A way of creating a copy of the given repository so that it belongs to our user

**Functions:** A reusable block of code that performs a specific task

**Futures:** A module provides a couple of different executors, one for using threads and the other one for using processes

## G

**Garbage collector:** A tool in charge of freeing the memory that's no longer in use

**Git:** A free open source version control system available for installation on Unix based platforms, Windows and macOS

**GitHub:** A web-based Git repository hosting service, allowing users to share and access repositories on the web and copy or clone them to a local computer

**Git directory:** A database for a Git project that stores the changes and the change history

**Git log:** A log that displays commit messages

**Git staging area:** A file maintained by Git that contains all the information about what files and changes are going to go into the next commit

**Globs:** Characters that create list of files, like the star and question mark

**grep:** An especially easy to use yet extremely powerful tool for applying RegExes

## H

**Head:** This points to the top of the branch that is being used

**Hot data:** Accessed frequently and stored in hot storage

**Hybrid cloud:** A mixture of both public and private clouds

## I

**IDE:** A software application that provides comprehensive facilities for software development

**Implicit conversion:** This occurs when the Python interpreter automatically converts one data type to another

**Incident commander** (incident controller): The person who needs to look at the big picture and decide what's the best use of the available resources

**Indirect merges:** GitHub can merge a pull request automatically if the head branch is directly or indirectly merged into the base branch externally

**Infinite loop:** A sequence that is missing a method for exiting the loop, causing the loop to run forever

**Infrastructure as a Service (or IaaS):** When a Cloud provider supplies only the bare-bones computing experience

**Input:** Information that is provided to a program by the end user

**Input / Output (I/O):** These streams are the basic mechanism for performing input and output operations in your programs

**Input/Output Operations Per Second (IOPS):** Measures how many reads or writes you can do in one second, no matter how much data you're accessing

**Integrated Development Environment (IDE):** A code editor with extra capabilities to simplify script writing

**Interpreter:** The program that reads and executes code

**Interpreter or interpreted language:** An intermediary program used to execute the instructions specified in the code

**Issue tracker (bug tracker):** A tracker that shows tasks that need to be done, the state they're in and who's working on them

**Iterators**: Variables that allow you to loop through a collection one item at a time

## J

**JSON:** A data-interchange format used in RESTful APIs to facilitate communication between clients and servers

## K

**Kernel:** The main core of an operating system that talks directly to hardware and manages the system’s resource

**Kubernetes:** An open-source platform that gives programmers the power to orchestrate containers

## L

**Latency:** The delay between sending a byte of data from one point and receiving it on the other

**Linear search:** The process of searching each line of data until the desired data entry is located

**Linux:** An open source operating system where the software is free to share, modify, and distribute

**Lists:** Sequences of elements

**List comprehensions:** Create new lists based on sequences or ranges

**Load balancer:** Ensures that each node receives a balanced number of requests

**Log files:** Log files are records or text files that store a history of events, actions, or errors generated by a computer system, software, or application for diagnostic, troubleshooting, or auditing purposes

**Logic errors:** Errors in code that prevent it from running correctly

**Logical operators:** Operators used to combine or manipulate boolean values (True or False) to create complex conditions for decision-making

**Lookahead:** RegEx that matches a pattern only if it’s followed by another pattern

**Loop:** A sequence that makes the computer do repetitive tasks

## M

**Machine language:** Lowest-level computer language. It communicates directly with computing machines in binary code (ones and zeros)

**Mac OS:** Operating system developed by Apple

**Manual scaling:** Changes are controlled by humans instead of software

**Master:** The default branch that Git creates for when a new repository initialized, commonly used to place the approved pieces of a project

**Memory leak:** This happens when a chunk of memory that's no longer needed is not released

**Memory profiler:** A tool used to figure out how the memory is being used

**Merge:** An operation that merges the origin/master branch into a local master branch

**Merge commits:** All commits from the feature branch are added to the base branch

**Merge conflict:** This occurs when the changes are made on the same part of the same file, and Git won't know how to merge those changes

**Modified files:** A stage where changes have been made to a file, but the have not been stored or committed

**Mode:** The format controlling what you can do with a recently opened file

**Multi-cloud:** A mixture of public and/or private clouds across vendors

## N

**NALSD (Non-Abstract Large System Design):** A discipline and process introduced by Google, primarily aimed at empowering site reliability engineers (SREs) to assess, design, and evaluate large-scale systems

**Naming conventions**: Functions, classes and methods with naming conventions to understand what to expect from them

## O

**Object storage:** Storage where objects are placed and retrieved into a storage bucket

**Observer effect:** The idea that observing a phenomenon alters the phenomenon

**Object-oriented programming language:** Most coding elements are considered to be objects with configurable properties

**Operating system (OS):** Software that manages everything that goes on in the computer, composed of two main parts: the kernel and the user space

**Orchestration:** The automated configuration and coordination of complex IT systems and services

**Output:** the end result of a task performed by a function or computer program

## P

**Parameter (argument):** A value passed into a function for use within the function, controlling the behavior of the CSV reader and writer

**Pass:** A placeholder statement which is used when the syntax requires a statement, but you don't want to execute any code or command

**Patch:** A command that can detect that there were changes made to the file and will do its best to apply the changes

**Persistent storage:** Storage used for instances that are long lived and need to keep data across reboots and upgrades

**Pipes:** A process of connecting the output of one program to the input of another

**Piping:** A process of connecting multiple scripts, commands, or other programs together into a data processing pipeline

**Pipelines:** The specific steps that need to run to obtain the desired result

**Platform as a Service (or PaaS):** When a Cloud provider offers a preconfigured platform to the customer

**Platform-specific scripting language:** Language used by system administrators on those specific platforms

**Private key:** A secret and secure cryptographic key that must be kept confidential and protected and is used to decrypt data that has been encrypted with the corresponding public key

**Programming:** The process of writing a program to behave in different ways

**Programming code**: A set of written computer instructions, guided by rules, using a computer programming language

**Programming languages:** Language with syntax and semantics to write computer programs

**Pod:** A group of one or more containers that are scheduled and run together

**Pointers:** The variables that store memory addresses

**Postmortems:** Documents that describe details of incidents to learn from mistakes

**Private cloud:** When your company owns the services and the rest of your infrastructure

**Production:** The software is pushed out to the end users from a cloud server

**Profiler:** A tool that measures the resources the code is using to see how the memory is allocated and how the time is spent

**Public cloud:** The cloud services provided to you by a third party

**Public key**: A safety cryptographic structure frequently employed to establish secure communication through data encryption or to validate the authenticity of a digital signature

**Pull request:** A procedure where new code is examined before it is merged to create a branch or master branch

**Puppet:** The current industry standard for configuration management, also known as the client

**Puppet master:** Known as the Puppet server

**Python:** A general purpose programming language

**Python interpreter:** Program that reads and executes Python code by translating Python code into computer instructions

**Pytest:** A powerful Python testing tool that assists programmers in writing more effective and stable programs

## Q

**Qwiklabs:** An online learning environment or virtual machine to simulate real-world scenarios

## R

**Rate limits:** Prevent one service from overloading the whole system

**Reader objects:**  Object that represents an element or entity within a scene that needs to be rendered to the screen

**Rebasing:** The base commit that's used for a branch is changed

**Real time:** The amount of actual time that it took to execute the command

**Recursion:** The repeated application of the same procedure to a smaller problem

**Redirection:** A process of sending a stream to a different destination

**Refactoring:** When a code is updated to be more self-documenting and clarify the intent

**Reference images:** Store the contents of a machine in a reusable format

**Registry:** A place where containers or artifacts are stored and organized

**Regular expression:** A search query for text that's expressed by string pattern, also known as RegEx or RegExp

**Relative path:** A portion of a path to show where the resource is located in relation to the current working directory

**Remote branches:** Git uses read-only branches to keep copies of the data that's stored in the remote repository

**Remote repositories:** Repositories that allow developers to contribute to a project from their own workstations making changes to local copies of the project independently of one another

**Repository:** An organization system of files that contain separate software projects

**Reproduction case:** A clear description of how and when the problem appears, a way to verify if the problem is present or not

**Resource Monitor** (or Performance Monitor): Windows OS tool that shows what's using the most CPU, memory, energy, disk, or network

**REST (Representational State Transfer):** Every request carries all the parameters and data needed for the server to satisfy that request

**RESTful APIs:** Rely on the HTTP protocol, can be further secured using HTTPS, and API endpoints can authenticate users via authorization tokens, API keys, or other security mechanisms

**REST architecture:** An architectural style for designing networked applications and web services

**Return value**: This is the value or variable returned as the end result of a function

**Richardson Maturity Model (RMM):** A framework that categorizes and describes different levels of implementation for RESTful APIs based on their adherence to the six constraints

**Rollback:** The act of reverting changes made to software to a previous state

## S

**Script:** Often used to automate specific tasks

**Secure Shell (SSH):** A robust protocol for connecting to servers remotely

**Semantics:** The intended meaning or effect of statements, or collections of words, in both human and computer languages

**Service-level agreements (SLAs):** An agreement between a vendor and its clients or users that can be legally binding

**Service-level objectives (SLOs):**  A specific and measurable target that defines the level of performance, reliability, or quality a service should consistently deliver

**Shell:** The application that reads and executes all commands

**Signals:** Tokens delivered to running processes to indicate a desired action

**Source Control Management (SCM):** A tool similar to VCS to store source code

**Software as a Service (or SaaS):** When a Cloud provider delivers an entire application or program to the customer

**Software testing:** A process of evaluating computer code to determine whether or not it does what is expected

**Standard input stream commonly (STDIN):** A channel between a program and a source of input

**Standard output stream (STDOUT):** A pathway between a program and a target of output, like a display

**Standard error (STDERR):** This displays output like standard out, but is used specifically as a channel to show error messages and diagnostics from the program

**Stage files:** A stage where the changes to files are ready to be committed

**Staging:** A strategic DevOps approach where we specify the build steps and tests

**Sticky sessions:** All requests from the same client always go to the same backend server

**String:** A data type used to represent a piece of text. sequences of characters and are immutable

**SSH client:** This establishes a connection to the SSH server, ensuring a secure interaction, where the client makes access requests

**SSH key:** An access credential

**SSH protocol:** Standard commonly used for logging in to servers remotely on the principle of public-key encryption

**SSH server:** This establishes secure network connections, undergoes mutual authentication, and initiates encrypted login sessions or file transfers

**Squash commits:** The decision add commit messages together and an editor opens to make any necessary changes

**Subprocesses:** A process to call and run other applications from within Python, including other Python scripts

**Swap:** A space in the hard drive where the operating system puts the parts of the memory that aren't currently in use

**Syntax:** The rules for how each statements are constructed in both human and computer languages

**Sys time:** The time spent doing system level operations

**System calls:** The calls that the programs running on our computer make to the running kernel

## T

**Technical debt:** The pending work that accumulates when a quick-and-easy solution is applied instead of a sustainable long-term one

**Templating:** The process of capturing all of the system configuration to let us create VMs in a repeatable way

**Test case:** This is the individual unit of testing that looks for a specific response to a set of inputs

**Test fixture:** This prepared to perform one or more tests

**Test suite:** This is used to compile tests that should be executed together

**Test runner:** This runs the test and provides developers with the outcome’s data

**Threads:** Run parallel tasks inside a process

**Three-way merge:** A merge when the snapshots at the two branch tips with the most recent common ancestor, the commit before the divergence

**Throughput:** The amount of data that you can read and write in a given amount of time

**Tracked:** A file’s changes are recorded

**Traffic shaping:** This is a way of marking the data packets sent over the network with different priorities, to avoid having huge chunks of data use all of the bandwidth

**Troubleshooting:** The process of solving any kind of problem in the system running the application

**Tuples:** Sequences of elements of any type that are immutable, written parentheses instead of square brackets

## U

**Undefined behavior:** The code is doing something that's not valid in that programming language

**unittest:** A set of Python tools to construct and run unit tests

**Unit tests:** A test to verify that small isolated parts of a program work correctly

**Untracked:** A file’s changes are not recorded

**User space:** Everything outside of the kernel that users interact with directly

**User time:** The time spent doing operations in the user space

**Utilization limits:** Cap the total amount of a certain resource that you can provision

## V

**Valgrind:** A powerful tool that can tell if the code is doing any invalid operations, no matter if it crashes or not

**Variables:** These are used to temporarily store changeable values in programming code

**Version control systems (VCS):** A tool to safely test code before releasing it, allow multiple people collaborate on the same coding projects together, and stores the history of that code and configuration

**Virtual environment:** A tool that allows you to create isolated environments for your Python projects

**Virtual machine (VM):** A computer simulated through software

## W

**Watchdog:** This is another process that checks whether a program is running and, when it's not, starts the program again

**Web application:** An application that you interact with over HTTP

**While loop:** This is used when a segment of code needs to execute repeatedly while a condition is true

**White-box test:** A test where test creator knows how the code works and can write test cases that use the understanding to make sure it performs as expected

**Wildcard:** A character that can match more than one character

**Windows OS:** Operating system developed by Microsoft

**Wrapper:** A function or program that provides a compatibility layer between two functions or programs, so that they can work well together

**Writer objects:** The capability to write data to a CSV file